Report for 2001DC3921B: Analysis of Transformed Environmental Data with Detection Limits

There are no reported publications resulting from this project.

Report Follows:

Properties of the Power-Normal Distribution

Jade Freeman, U.S. Environmental Protection Agency
Reza Modarres, Department of Statistics, George Washington University
(Reza@GWU.Edu)

Abstract:

Box-Cox transformation system produces the power normal (PN) family, whose members include normal and log-normal distributions. We study the moments of PN and obtain expressions for its mean and variance. The quantile functions are discussed. The conditional distributions are studied and shown to belong to the PN family. We obtain expressions for the mean, median and modal regressions. Chevyshev-Hermite polynomials are used to obtain an expression for the correlation coefficient and to prove that correlation is smaller in the PN scale than the original scale. We use the Fréchet bounds to obtain expressions for the lower and upper bounds of the correlation coefficient.

Key Words: Box-Cox Transformation, Log-Normal, Uncertainty Analysis.

1. Introduction

When $\{X_i\}$ are independent and positive random variables Galton (1879) showed that the limiting distribution of $\prod_{i=1}^{n} X_i$ on the log-scale, i.e., $\sum_{i=1}^{n} \log X_i$, is normal as n approaches infinity. The distribution of this product in the original scale is well approximated with a two-parameter log-normal distribution. The result is exact when the $\{X_i\}$ are log-normal. More generally, one may consider a power transformation, rather than logarithm, of an underlying normal process. Consider the Box-Cox (1964) power transformation: $Y = P_\lambda(X) = (X^\lambda - 1)/\lambda$ when $\lambda \neq 0$ or $Y = \ln(X)$ when $\lambda = 0$. Frequently, Y is approximated with a normal distribution with mean μ and variance σ^2. After a Box-Cox transformation, one is often interested in inference on the original scale; e.g. estimate the mean (Shumway, Azari and Johnson, 1989; Freeman and Modarres, 2002a), requiring a back transformation. This is often difficult as the mean is a non-linear function of μ and σ^2 (Land, 1974). An equivalent strategy is to model the data directly in the original scale in terms of functions of normal variates. Box-Cox transformation has been successful in many applications and the subject of numerous investigations (Sakia, 1992). Whenever Box-Cox transformation is effective, one may argue that the observations in the original scale must be well approximated by powers of normal variates. It is the purpose of this article to study the family of distributions obtained through this system.

The analysis of environmental data frequently centers on positive random variables such as the concentration of pollutants. Such concentrations are usually right skewed with several extreme observations at both low and high levels. A parametric model such as log-normal, gamma, Weibull or Inverse Gaussian (Haas, 1997; Ott, 1995) is often used to model the observations. However, we often do not have adequate knowledge (e.g. sample size) to specify a distributional form; i.e. a clear fit is not obtained through goodness of fit tests. Hence, model uncertainties exist. In such cases transformation to normality, or equivalently, analysis on the PN scale is an appealing alternative. One can study model uncertainties through the transformation parameter of a PN distribution. Frequently, the log-normal model is selected based on chemical, biological, or physical grounds (Ott, 1995) and it has a prominent role in many application areas (Johnson, Kotz, and Balakrishnan, 1994).

Much effort has been exerted to research the Box-Cox transformation. With the exception of the work of Goto and Inoue (1980), relatively little is known about the distribution of the variables in the PN scale. Much of the available results pertain to the log-normal distribution. In the next section, we discuss the PN family, its moments and quantile function. We develop the multivariate PN distribution in section 3, where we study the conditional distributions and show that they are also in the PN family. We also investigate the mean, median and modal re-

The work of Reza Modarres was supported by a cooperative agreement with the Office of Water of the USEPA. The work of Jade Freeman was supported by a grant from the USGS through the Water Resources Research Institute of the University of District of Columbia.

gressions of this family. Chevyshev-Hermite polynomials are used in section 4 to derive an expression for the correlation coefficient and to prove it is smaller in the PN scale than the original scale. We use the Fréchet bounds to obtain expressions for the lower and upper bounds of the correlation coefficient.

2. Power Normal Distribution

Johnson (1949) considers a transformation system which includes normal, lognormal, $\sinh^{-1}$-normal, and logit-normal. Johnson (1987) uses the system for generating variates for statistical simulation and further develops it to a multivariate system. Here, we discuss the Box-Cox power transformation system and the PN distribution. By applying an inverse transformation to a normal random variable Y, one obtains $BC(\lambda) : X = (\lambda Y + 1)^{1/\lambda}$ for $\lambda \neq 0$ and $X = \exp(Y)$ for $\lambda = 0$. The system produces the PN family of distributions. This family was first noted in Goto and Inoue (1980), where the authors investigate some of its properties. We discuss other aspects of this family in this article and concentrate on $0 \leq \lambda \leq 1$, which includes several well-known transformations such logarithm, square, cube or fourth roots (see Shumway et al, 1989).

Researchers have generally assumed that there is a transformation parameter that produces a normal distribution for all λ. Since the support of X is positive, Y has a truncated normal distribution for $\lambda \neq 0$. Let $Y \sim TN(\mu, \sigma^2, -1/\lambda)$ have a truncated normal density function $g(Y \mid \mu, \sigma^2, -1/\lambda) = \frac{1}{K}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}(Y-\mu)^2\right\}$ where $K = \Phi(T)$, 1, or $\Phi(-T)$ when $\lambda > 0$, $\lambda = 0$ or $\lambda < 0$, respectively. Note that $T = (1/(\lambda\sigma) + 1/\kappa)$ where $\kappa = \sigma/\mu$ is the coefficient of variation and K is a normalizing constant that corresponds to the area above or below the point of truncation, $-1/\lambda$. Let $X \sim PN(\lambda, \mu, \sigma^2)$ denote a PN random variable with pdf, for $X > 0$, $f(X \mid \lambda, \mu, \sigma^2) = \frac{1}{K}\frac{1}{\sqrt{2\pi}\sigma}X^{\lambda-1} \cdot \exp[-\frac{1}{2}(\frac{p_\lambda(X)-\mu}{\sigma})^2]$. By differentiating the density function, one can show that the distribution is unimodal in the interval $0 \leq \lambda \leq 1$ as κ approaches zero. Let $\delta = (1+\lambda\mu)^2 + 4\sigma^2\lambda(\lambda-1)$. The density has a mode at

$$Mode(X) = \begin{cases} [0.5(1+\lambda\mu+\sqrt{\delta})]^{1/\lambda}, & \lambda \neq 0, \\ \exp(\mu - \sigma^2), & \lambda = 0. \end{cases} \tag{1}$$

The density is right skewed for $0 \leq \lambda < 1$. For $\lambda > 0$, as $\kappa \to 0$, the standardized point of truncation $T \to \infty$ and the left tail of Y is no longer truncated.

2.1 Moments

One can show that the rth moment in the PN scale is a non-linear function of the means and variances in the transformed scale. When $\lambda > 0$, we have

$$E(X^r) = \int_{-1/\lambda}^{\infty} (\lambda y + 1)^{r/\lambda}\phi\left(\frac{y-\mu}{\sigma}\right)\frac{dy}{d\sigma}. \tag{2}$$

In this section, we obtain several more useful forms for the moments of a PN distribution. For $\lambda = 0$, note that $E(X^r) = E(\exp(rY)) = \exp(r\mu + r^2\sigma^2/2)$ and for $\lambda > 0$, we have $Y \sim TN(\mu, \sigma^2, -1/\lambda)$. Let $S(y) = (\lambda y + 1)^{r/\lambda}$. One can expand $S(y)$ around μ to show $S(y) = \sum_{i=0}^{\infty}\frac{1}{i!}S^{(i)}(\mu)(y-\mu)^i$ where $S^{(i)}(y) = (\lambda y+1)^{r/\lambda - i}\prod_{j=1}^{i-1}(r - j\lambda)$ and obtain

Lemma 1 : Let $X \sim PN(\lambda, \mu, \sigma^2)$. If Y=$(X^\lambda - 1)/\lambda$ and $Z = (Y-\mu)/\sigma$. One has

$$E(X^r) = \sum_{i=0}^{\infty}\frac{1}{i!}S^{(i)}(y)\sigma^i E(Z^i), \tag{3}$$

for $\lambda > 0$ and $E(X^r) = \exp(r\mu + \frac{r^2\sigma^2}{2})$ for $\lambda = 0$.

Note that $Z \sim TN(0,1,T)$ and that $E(Z^i) = \frac{\phi(T)}{1-\Phi(T)}H_{i-1}(T) + R$ where R is a polynomial of degree $i-2$ in Z and H_{i-1} is the $(i-1)th$ Chevyshev-Hermite Polynomial. When Y is approximated with a normal distribution; e.g. for small κ, we have $E(y-\mu)^i = (\sigma^i i!)/(2^{i/2}(i/2)!)$ for even i and 0 for odd i. Therefore,

Lemma 2 : Let $X \sim PN(\lambda, \mu, \sigma^2)$, $\lambda \neq 0$ and $Y \sim N(\mu, \sigma^2)$. Then,

$$E(X^r) = \sum_{Even\ i \geq 0}\frac{\sigma^i i!}{2^{i/2}(i/2)!}S^{(i)}(y). \tag{4}$$

Tables 1 and 2 obtain the form of $E(X)$ and $Var(X)$ for some $0 \leq \lambda \leq 1$. Let $\delta_i = (\lambda\sigma)^i(\lambda\mu+1)^{(m-i)}$. When $m = r/\lambda$ is an integer, one can show for $\lambda \neq 0$, that $E(X^r)$ is

$$\begin{cases} \sum_{i=0}^{m}\binom{m}{i}\delta_i E(Z^i), \\ \sum_{\text{Even } i=0}^{m}\binom{m}{i}\delta_i i!/(2^{i/2}(i/2)!) \end{cases} \tag{5}$$

where $Y \sim TN(\mu, \sigma^2, -1/\lambda)$ and $Y \sim N(\mu, \sigma^2)$, respectively. For example, when $r = \lambda$, $E(X^\lambda) = \lambda\mu + 1$ and $Var(X^\lambda) = \lambda^2\sigma^2$.

2.2 CDF and the Quantile Functions

One can consider median and other quantiles to avoid difficulties with the mean. CDF and quantile functions can be used as tools in statistical modeling in a number of applications when interest focuses particularly on the extreme observations in the tails of the data (Modarres, Nayak and Gastwirth, 2002). For example, to identify a suitable model, graphs and exploratory analysis of sample observations will give an impression of the basic shape of distribution. Adequacy of fit can be judged from a plot of sample quantiles against the corresponding model quantiles. Let $Z = (p_\lambda(X) - \mu)/\sigma$. The cdf of $PN(\lambda, \mu, \sigma^2)$ is $F(X) = \frac{1}{K} \cdot (\Phi(Z) - \Phi(-T))$ for $\lambda > 0$, and $\frac{1}{K} \cdot \Phi(Z)$ for $\lambda < 0$. When $\lambda = 0$ or $Y \sim N(\mu, \sigma^2)$ one has $F(X) = \Phi(Z)$. For large T, $F(X) = \Phi(p_\lambda(X) - \mu)/\sigma)$. Let $V(p) = 1-(1-p)\Phi(T)$ for $0 < p < 1$. The quantile function of $PN(\lambda, \mu, \sigma^2)$ is given by $Q_\lambda(p) =$

$$\begin{cases} (\lambda(\sigma\Phi^{-1}(V(p)) + \mu) + 1)^{1/\lambda}, & \lambda > 0, \\ \exp(\mu + \sigma\Phi^{-1}(p)), & \lambda = 0, \\ (\lambda(\sigma\Phi^{-1}(p) + \mu) + 1)^{1/\lambda}, & \lambda < 0. \end{cases} \tag{6}$$

One can obtain a simultaneous quantile plot for different values of λ. Such a plot reveals that the transformation parameter λ has more effects on the upper tail for $0 \leq \lambda \leq 1$ and that extreme observations may have more influence on estimation of λ. The log-normal quantile function has a longer tail and is clearly separated from other PN quantiles. This explains why likelihood-based methods of model selection perform so well in identifying the lognormal distribution (Shumway et al., 1989). One may obtain a weighted least squares estimator for λ by modifying the least squares estimator of λ. Maximum likelihood estimation of the quantiles is an attractive procedure due to the existing form of (6). One can use the asymptotic normality of MLE's along with their invariance property to show asymptotic normality of $\hat{Q}_\lambda(p)$ using $\hat{\mu}$ and $\hat{\sigma}^2$, which are MLE's of the mean and variance on the transformed scale.

3. Multivariate Power-Normal

Consider the Box-Cox power transformation defined by $p_{\lambda_j}(X_j) = \frac{X_j^{\lambda_j}-1}{\lambda_j}$ when $\lambda_j \neq 0$ and $p_{\lambda_j}(X_j) = \ln X_j$ when $\lambda_j = 0$ for each variable X_j, $j = 1, ..., p$, that is non-negative. Let $Q = [p_{\vec{\lambda}}(\vec{X}) - \vec{\mu}]'\Sigma^{-1}[p_{\vec{\lambda}}(\vec{X}) - \vec{\mu}]$. The inverse transformations define a p-variate vector $\vec{X} = (X_1, X_2, ..., X_p)$ with a probability distribution $f(\vec{X} \mid \vec{\lambda}, \vec{\mu}, \Sigma) =$

$$\frac{1}{K} \cdot \frac{1}{(2\pi)^{p/2} \mid \Sigma \mid^{1/2}} \prod_{j=1}^{p} X_j^{\lambda_j - 1} \cdot \exp(-\frac{1}{2}Q),$$

where K depends on $T_i = 1/(\lambda_i\sigma_i) + 1/\kappa_i$. Denote the bivariate standard normal pdf and cdf with $\phi_2(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp[-\frac{1}{2(1-\rho^2)}(z_1^2 + z_2^2 - 2\rho z_1 z_2)]$ and $\Phi_2(z_1, z_2) = \int_{-\infty}^{z_1}\int_{-\infty}^{z_2} \phi_2(t_1, t_2)dt_1dt_2$, respectively. The random vector (X_1, X_2) from $PN(\vec{\lambda}, \vec{\mu}, \Sigma)$ has a pdf $f(X_1, X_2) = \frac{1}{K} \cdot \frac{1}{\sigma_1\sigma_2} X_1^{\lambda_1 - 1} X_2^{\lambda_2 - 1} \phi_2(Z_1, Z_2)$ for $X_i > 0$, i= 1, 2, where $\vec{\lambda} = (\lambda_1, \lambda_2)$, $\vec{\mu} = (\mu_1, \mu_2)$, $\Sigma = (\sigma_{ij})$. Note that $K = \Phi_2[S(\lambda_i)T_i, S(\lambda_j)T_j]$ for $\lambda_i \neq 0$ and $\lambda_j \neq 0$, and $K = \Phi(S(T_i)T_i)$ for $\lambda_i \neq 0$ and $\lambda_j = 0$ where $S(\lambda_i)$ refers to the sign of the transformation, For the bivariate log-normal distribution, $K = 1$ when $\lambda_i = \lambda_j = 0$.

Assuming that the joint distribution of $\vec{Y} = p_{\vec{\lambda}}(\vec{X})$ is approximately normal, the forms for the covariance of the selected bivariate PN distributions are given in Table 3. The distribution of (Y_1, Y_2) is truncated bivariate normal for $\lambda_i \neq 0$, i=1, 2, and bivariate normal for (λ_1, λ_2)=(0,0). As in the univariate case, (Y_1, Y_2) are approximately bivariate normal for small coefficient of variations κ_1 and κ_2 for $\vec{\lambda} \neq \vec{0}$. Most researchers assume $K \approx 1$ for practical purposes (Johnson and Wichern, 2002; Gnanadesikan, 1977).

As an aid for model selection, a collection of contours and 3-dimensional plots for several values of $\vec{\lambda}$ appear in Freeman and Modarres (2002b). These plots are helpful in the early stages of model selection. Examination of bivariate contour plots along with univariate Q-Q plots help to identify a transformation set. The likelihood function can be maximized over this set to obtain an effective scale on which to analyze the data. It is interesting to note that the bivariate contours change forms as the forms of the margins change and the elliptical shape of the bivariate contours vanish when the margins are not normal even though the joint dependence is through a bivariate normal copula. One can show that if (y_1, y_2) has a bivariate normal distribution, then $F(X_1, X_2) = \Phi_2(Z_1, Z_2)$. The next lemma, which follows from properties of a multivariate normal distribution (Anderson, 1984), states that the conditional distributions derived from joint PN distributions are also PN.

Lemma 3 : Let $\vec{X}$ be the PN p-variate random vector such that $\vec{X} = \left(\vec{X}^{(1)}, \vec{X}^{(2)}\right)$ where $\vec{X}^{(1)}$ and $\vec{X}^{(2)}$ are q and $(p-q)$-variate random vectors with parameter vectors $\vec{\lambda}^{(1)}$ and $\vec{\lambda}^{(2)}$. Suppose that we partition $\vec{\mu}$ and Σ similarly.

- *The marginal distribution of $\vec{X}^{(1)}$ is $PN(\vec{\lambda}^{(1)}, \vec{\mu}^{(1)}, \Sigma_1)$.*
- *The conditional distribution of $\vec{X}^{(2)}$ given $\vec{X}^{(1)}=\vec{x}^{(1)}$ is a $(p-q)$-variate $PN(\vec{\lambda}^{(2)}, \vec{\mu}^*, \Sigma^*)$ with $\vec{\mu}^* = \vec{\mu}^{(2)} + \Sigma_{12}\Sigma_{11}^{-1}(p_{\vec{\lambda}^{(1)}}(\vec{X}^{(1)}) - \vec{\mu}^{(1)})$ and $\Sigma^* = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.*
- *$\vec{X}^{(1)}$ and $\vec{X}^{(2)}$ are independent when $\Sigma_{12} = 0$.*

Note that the conditional means depend non-linearly and the variances and covariances do not depend on the values of the fixed variates. For example, the conditional distribution of X_2 given $X_1 = x_1$ is $PN(\lambda_2, \mu, \sigma^2)$ where $\mu = \mu_2 + \rho\sigma_2/\sigma_1(p_{\lambda_1}(x_1) - \mu_1)$ and $\sigma^2 = \sigma_2^2(1-\rho^2)$. Mostafa and Mahmoud (1964) study the mean, median and modal regression of the bivariate log-normal distribution. We extend their result to the PN distribution in the following lemma.

Lemma 4 : If (X_1, X_2) has a bivariate PN distribution, then Median($X_2|X_1$), $E(X_2|X_1)$, and Mode($X_2|X_1$) are obtained by evaluating (1), (3) and (6), respectively, at $\mu = \mu_2 + \rho\sigma_2/\sigma_1(p_{\lambda_1}(x_1) - \mu_1)$ and $\sigma^2 = \sigma_2^2(1-\rho^2)$ with $\lambda = \lambda_2$.

4. Correlations

Let ρ_{X_1,X_2} and $\rho = \rho_{Y_1,Y_2}$ denote the coefficient of correlations in the PN and normal scales, respectively. In this section we assume $\vec{Y} \sim (\vec{\mu}, \Sigma)$ and show that $\rho_{X_1,X_2} = f(\rho)$, where the form of the function f depends on the transformation parameter $\vec{\lambda}$ as well as $\vec{\mu}$, and Σ. See (Freeman and Modarres, 2002b) for a more complete discussion. One can show

$$f(\rho) = \frac{\sum_{i=1}^{\infty} b_{1i}b_{2i}\rho^i i!}{\sqrt{(\sum_{i=1}^{\infty} b_{1i}^2 i!)(\sum_{j=1}^{\infty} b_{2j}^2 j!)}}. \tag{7}$$

It follows from the form of the joint density of (X_1, X_2) and equation (7) that if $\rho = 0$, then $f(\rho) = 0$. Further, ρ is zero when $f(\rho) = 0$ by the transformation property of functions of independent random variables (Karr, 1993). The following lemma whose proof appears in the technical report shows that the coefficient of correlation in the PN scale cannot be greater than that in the normal scale.

Lemma 5 : Let (X_1, X_2) be distributed with $PN(\vec{\lambda}, \vec{\mu}, \Sigma)$ with $\vec{\mu} = (\mu_1, \mu_2)$, and covariance Σ. Let $f(\rho)$ be the correlation of coefficient of X_1 and X_2. Then, $\mid f(\rho) \mid \leq \mid \rho \mid$.

For the bivariate log-normal distribution this result is given without proof in Mostafa and Mahmoud (1964). Finally, note that $P_\lambda(X)$ are monotone transformations and the rank measures of correlation remain the same on both scales.

4.1 Extreme Correlations

Even though the minimum and maximum correlation for $\vec{\lambda} = (0, 0)$ and $\vec{\lambda} = (1, 1)$ are -1 and 1, it may not be the case for other transformations. Specific forms of extreme correlations of selected (λ_1, λ_2) are obtained and appear in Table 3. It is tedious to determine the mathematical forms of extreme correlations. One can, however, use the following computational scheme to calculate them numerically. Let Π be the set of all cdf's $F(X_1, X_2)$ on R^2 having marginal cdf's $F_i(X_i)$, $i = 1, 2$, with finite variances. Fréchet's bounds (Fréchet, 1951) provide $H_0(X_1, X_2) \leq F(X_1, X_2) \leq H_1(X_1, X_2)$ where $H_0(X_1, X_2) = Max(0, F_1(X_1) + F_2(X_2) - 1))$ and $H_1(X_1, X_2) = Min(F_1(X_1), F_2(X_2))$ belong to Π.

To show that the correlations under H_0 and H_1 are the minimum and maximum, respectively, note that $f(\rho) = \frac{1}{\sigma_1\sigma_2}(\int_0^\infty \int_0^\infty F(x_1, x_2) - F_1(x_1)F_2(x_2)dx_1dx_2)$ (Lehmann, 1966). Let ρ_0 and ρ_1 be the coefficient of correlation under H_0 and H_1. It follows that $\rho_0 \leq f(\rho) \leq \rho_1$. Let Q_j be the quantile function of X_j for $j = 1, 2$ and u be a uniform random variable in the intervel $(0, 1)$. One can show (See Whitt, 1976) that $(Q_1(u),\ Q_2(1-u))$ has cdf $H_0(X_1, X_2)$ and $(Q_1(u),\ Q_2(u))$ has cdf $H_1(X_1, X_2)$. It follows that $Corr[(Q_1(u), Q_2(1-u))] \leq f(\rho) \leq Corr(Q_1(u), Q_2(u))$.

To obtain the numerical values for the minimum and maximum correlations of $PN(\vec{\lambda}, \vec{\mu}, \Sigma)$, we generate a vector of independent uniform random variables $\vec{u}_n$; then, the maximum and minimum correlations for $(\lambda_1 > 0, \lambda_2 > 0)$ are computed from $f^{min}(\rho) = Corr(Q_1(u), Q_2(1-u))$ and $f^{max}(\rho) = Corr(Q_1(u), Q_2(u))$ where the quantile function is given by (6). Technical report contains a table of the averages of 100 minimum and maximum correlations computed based on $n = 100,000$, $\vec{\mu}$=(4,4), and $\sigma_1^2 = \sigma_2^2 = 1$ for each selected $\vec{\lambda}$.

References

[1] Anderson, T. W. (1984). An Introduction to Multivariate Analysis. Second Edition. Wiely.

[2] Box, G. E., and Cox, D. R. (1964). An Analysis of Transformed Data. Journal of Royal Statistical Society. B, 39, 211-252.

[3] David, F. N. and Johnson, N. L. (1954). A Test for Skewness with Ordered Variables. Annals of Eugenics. 18, 351-353.

[4] Fréchet, M. (1951). Sur Les Tableaux de Corrélation Dont le Marges sont Données. Ann. Univ: Lyon Sect. A 14, 53-77.

[5] Freeman, J., and Modarres, R. (2002a). Analysis of Censored Environmental Data with Box-Cox Transformations. Technical Report. Department of Statistics. George Washington University.

[6] Freeman, J., and Modarres, R. (2002b). Properties of the Power Normal Distribution. Technical Report. Department of Statistics. George Washington University.

[7] Galton, F. (1879). The Geometric Mean in Vital and Social Statistics. Proceedings of the Royal Statistical Society of London. Volume 29, 365-367.

[8] Gnanadesikan, R. (1977). Methods for Statistical Analysis of Multivariate Observations. John Wiley & Sons. New York.

[9] Goto, M. and Inoue, T. (1980). Some Properties of Power Normal Distribution. Bulletin of the Biometric Society of Japan, 1, 28-54.

[10] Haas, C. N. (1997). Importance of Distributional Form in Characterizing Inputs to Monte Carlo Risk Assessment. Risk Analysis. Volume 17, No. 1, 107-113.

[11] Johnson, M. E. (1987). Multivariate Statistical Simulation, New York.

[12] Johnson, N. L. (1949). Systems of Frequency Curves Generated by Methods of Translation. Biometrika, Vol. 36, No. 1/2, 149-176.

[13] Johnson, R. A., and Wichern, D. W. (2002). Applied Multivariate Analysis. Fifth Edition, Prentice Hall.

[14] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). Continuous Univariate distributions. Volumes I. John Wiley & Sons, 2nd ed.

[15] Karr, A. F. (1993). Probability. Springer, New York.

[16] Land, C. E. (1974). Confidence Interval Estimation for Means After Data Transformations to Normality. Journal of the American Statistical Association, Vol. 69, No. 347, 795-802.

[17] Lehmann, E. L. (1966). Some Concept of Dependence. Annals of Mathematical Statistics. 37, 1137-1153.

[18] Modarres, R., Nayak, T. K. and Gastwirth, J. L. (2002). Estimation of Upper Quantiles under Model and Parameter Uncertainty. Computational Statistics and Data Analysis, 39, 529-554.

[19] Mostafa, M. D. and Mahmoud, M. W. (1964). On the Problem of Estimation for the Bivariate Lognormal Distribution. Biometrika, Vol. 51, No. 3/4. 522-527.

[20] Ott, W. R. (1995). Environmental Statistics and Data Analysis. Lewis Publishers.

[21] Sakia, R. M. (1992). Box-Cox Transformation Technique: A Review. The Statistician. 42, 169-178.

[22] Shumway, R. H., Azari, A.S., and Johnson, P. (1989). Transformation for Environmental Data with Detection Limits. Technometrics. Volume 31, No. 3, 347-356.

[23] Whitt, W. (1976). Bivariate Distributions with Given Marginals. Annals of Statistics. 4, 6, 1280-1289.

Table 1: Means of Power-Normal Distribution

λ	E(X)
0	$\exp(\mu + \frac{1}{2}\sigma^2$
$\frac{1}{4}$	$\frac{1}{4}\mu + 1)^4 + \frac{3}{8}\sigma^2(\frac{1}{4}\mu + 1)^2 + \frac{3}{256}\sigma^4$
$\frac{1}{3}$	$\left(\frac{1}{3}\mu + 1\right)^3 + \frac{1}{3}\sigma^2\left(\frac{1}{3}\mu + 1\right)$
$\frac{1}{2}$	$\frac{1}{2}\mu + 1)^2 + \frac{1}{4}\sigma^2$
1	$\mu + 1$

Table 2: Variances of Power-Normal Distribution

λ	Var(X)
0	$\exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$
$\frac{1}{4}$	$\frac{8}{2048}\sigma^8 + (\frac{1}{4}\mu + 1)^2\sigma^2(\frac{3}{32}\sigma^4 + \frac{21}{32}\sigma^2\left(\frac{1}{4}\mu + 1\right)^2 + \left(\frac{1}{4}\mu + 1\right)^4)$
$\frac{1}{3}$	$\frac{5}{243}\sigma^6 + \frac{4}{9}\sigma^4\left(\frac{1}{3}\mu + 1\right)^2 + \sigma^2\left(\frac{1}{3}\mu + 1\right)^4$
$\frac{1}{2}$	$\frac{1}{8}\sigma^4 + \sigma^2\left(\frac{1}{2}\mu + 1\right)^2$
1	σ^2

Table 3: Covariances of Bivariate Power-Normal Distribution

(λ_1, λ_2)	$\mathrm{Cov}(X_1, X_2)$
$(0, 0)$	$\exp(\mu_1 + \mu_2 + \frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2)(\exp(\rho\sigma_1\sigma_2) - 1)$
(0,1/2)	$\rho\sigma_1\sigma_2\exp(\mu_1 + \frac{1}{2}\sigma_1^2)(\frac{1}{4}\rho\sigma_1\sigma_2 + \frac{1}{2}\mu_2 + 1)$
(0,1)	$\rho\sigma_1\sigma_2\exp(\mu_1 + \frac{1}{2}\sigma_1^2)$
(1/2,1/2)	$\frac{1}{8}(\rho\sigma_1\sigma_2)^2 + \rho\sigma_1\sigma_2\frac{1}{2}(\frac{1}{2}\mu_1\mu_2 + \mu_1\mu_2 + 2)$
(1/2,1)	$\rho\sigma_1\sigma_2(\frac{1}{2}\mu_1 + 1)$
(1,1)	$\rho\sigma_1\sigma_2$